

QUASI-HOMOMORPHISMS ON MAPPING CLASS GROUPS

MLADEN BESTVINA AND KOJI FUJIWARA

University of Utah, USA and Tohoku University, Japan

To Sibe Mardešić, with admiration

ABSTRACT. We refine the construction of quasi-homomorphisms on mapping class groups. It is useful to know that there are unbounded quasi-homomorphisms which are bounded when restricted to particular subgroups since then one deduces that the mapping class group is not boundedly generated by these subgroups. In this note we enlarge the class of such subgroups. The generalization is motivated by considerations in first order theory of free groups.

1. INTRODUCTION

Recall that a *quasi-homomorphism* on a group G is a function $h : G \rightarrow \mathbb{R}$ such that

$$\sup_{\gamma_1, \gamma_2 \in G} |h(\gamma_1 \gamma_2) - h(\gamma_1) - h(\gamma_2)| < \infty.$$

The set of all quasi-homomorphisms on G forms a vector space $\mathcal{V}(G)$. Any bounded function on G is trivially a quasi-homomorphism, and we consider the vector space $QH(G)$ which is the quotient of $\mathcal{V}(G)$ by the subspace of bounded functions. The existence of many quasi-homomorphisms $G \rightarrow \mathbb{R}$ has implications for (stable) commutator length [Ba], (failure of) bounded generation of G , and non-embedding results for arithmetic lattices. For example, if $h : G \rightarrow \mathbb{R}$ is an unbounded quasi-homomorphism such that the restrictions

2000 *Mathematics Subject Classification.* 57M07, 57S30.

Key words and phrases. Mapping class group, quasi-homomorphism, bounded cohomology, pseudo-Anosov.

The first author gratefully acknowledges the support by the National Science Foundation.

$h|Q_i$ to subgroups Q_1, \dots, Q_k are bounded, then G is not boundedly generated by the Q_i 's, i.e. for no $N > 0$ can every element of G be written as the product of $\leq N$ elements of $Q_1 \cup \dots \cup Q_k$.

In this paper we consider the case when G is the mapping class group $MCG(S)$ of a compact surface S (this is the group of isotopy classes of homeomorphisms $S \rightarrow S$). With the exception of a small number of sporadic surfaces, $MCG(S)$ is a “large” group, e.g. it contains a nonabelian free subgroup. In [BeFu] we proved that $QH(MCG(S))$ is infinite-dimensional provided $MCG(S)$ is not virtually abelian (and we only considered the case of orientable S). More precisely, we have

THEOREM 1.1. *Let S be a compact surface such that $MCG(S)$ is not virtually abelian. Then $QH(MCG(S))$ is infinite-dimensional. Moreover, for every finite collection of cyclic subgroups*

$$C_1, \dots, C_k \subset MCG(S)$$

there is an infinite-dimensional subspace of $QH(MCG(S))$ such that each representative quasi-homomorphism h of any element in this subspace satisfies:

- (1) *h is bounded on each C_i , and*
- (2) *h is bounded on the stabilizer $H(\alpha)$ of every essential (non-degenerate if S is non-orientable, see Appendix) simple closed curve (or scc) α in S .*

Since we did not state the theorem in exactly this form in [BeFu] we will outline the proof in Section 3. Also, in [BeFu] we discussed only the case when S is orientable. We explain how to modify the argument when S is non-orientable.

COROLLARY 1.2. *Suppose $MCG(S)$ is not virtually abelian. Let*

$$Q_1, \dots, Q_m$$

be a finite collection of subgroups of $MCG(S)$ such that each Q_i is either cyclic or the stabilizer of an essential (non-degenerate) scc. Then

$$Q_1 Q_2 \cdots Q_m \neq MCG(S).$$

Indeed, any finite union of sets of the above form is a proper subset of $MCG(S)$.

Here the set on the left consists of compositions $\phi_1 \cdots \phi_m$ for $\phi_i \in Q_i$. This statement generalizes the fact that $MCG(S)$ is not boundedly generated (by cyclic subgroups) [FLM].

In this note we will extend Theorem 1.1 so that (2) includes a larger class of subgroups of $MCG(S)$. The additional subgroups will be called *D-subgroups* and they consist of mapping classes that descend down a particular proper covering map $p : S \rightarrow S'$. The motivation for seeking such an extension comes from the first order theory of free groups, which also demands a

consideration of non-orientable surfaces. In this theory one is interested in understanding the set of conjugacy classes of homomorphisms $f : H \rightarrow \mathbb{F}$ that extend to $\tilde{f} : G \rightarrow \mathbb{F}$, where \mathbb{F} is a fixed nonabelian free group, and $H \subset G$ are fixed finitely generated groups. For example, consider the special case when H is the fundamental group of a closed surface S . To glean the structure of this set it is useful to consider its intersections with “orbits”: Let $f : H \rightarrow \mathbb{F}$ be a fixed homomorphism and consider the set

$$E_f = \{\phi \in MCG(S) \mid f\phi \text{ extends to } G\}$$

With the help of Makanin-Razborov diagrams [Se] (see also [KM1, KM2]) one shows that either $E_f = MCG(S)$ or E_f is contained in a finite union of sets of the form

$$Q_1 Q_2 \cdots Q_m$$

where each Q_i is either cyclic, or it is contained in the stabilizer of a simple closed curve, or it is a D-subgroup. The details of this are in [BeFe], where a sharper statement is proved. Thus Corollary 1.2 implies (using our extension of Theorem 1.1) that E_f is either all of $MCG(S)$ or else it is a “slim” subset of $MCG(S)$.

This reasoning applies to groups H other than surface groups by considering abelian JSJ decompositions [DS, FuP]. The most interesting pieces of the decomposition are surfaces (with boundary) and we are led to considering non-closed surfaces as well. It is convenient to collapse boundary components to punctures. This does no harm, since it only means that the quasi-homomorphisms we construct will be 0 on the subgroup generated by Dehn twists in the boundary components.

The crucial tool in the study of mapping class groups is that of the *curve complex* associated to the surface. This complex was originally defined by Harvey [Ha], it was used by Harer [Har] in his study of the homology of mapping class groups, and its geometric aspects have been studied more recently thanks to the celebrated theorem of Masur and Minsky which states that the curve complex is δ -hyperbolic. We review the definition and basic properties of the curve complex in the Appendix, where we also outline proofs of basic facts for the case of non-orientable surfaces.

2. D-SUBGROUPS OF MAPPING CLASS GROUPS

Let $p : S \rightarrow S'$ be a covering map between two compact connected surfaces. We will also assume that p is not a homeomorphism. There is a finite index subgroup $MCG_p(S') < MCG(S')$ consisting of (isotopy classes of) homeomorphisms $\phi' : S' \rightarrow S'$ that admit a lift, i.e. a homeomorphism $\phi : S \rightarrow S$ such that $p\phi = \phi'p$. When a lift exists it may not be unique. Consider the group

$$D(p) = \{\phi : S \rightarrow S \mid p\phi = \phi'p \text{ for some } \phi' : S' \rightarrow S'\} / \text{isotopy} < MCG(S)$$

consisting of all possible lifts of elements of $MCG_p(S')$. Equivalently, this is the group of mapping classes in $MCG(S)$ that *descend* to S' via p .

The group $D(p)$ is *commensurable* with $MCG_p(S')$ in the sense that the group

$$\tilde{D}(p) = \{(\phi, \phi') \mid p\phi = \phi'p\} / \text{isotopy}$$

surjects to both with finite kernel.

Any subgroup of $MCG(S)$ of the form $D(p)$ will be called a *D-subgroup* of $MCG(S)$. Such subgroups arise naturally in the study of the first order theory of free groups and they play a role similar to the stabilizer of a simple closed curve in S .

LEMMA 2.1. *If $\chi(S) < 0$ then there are only finitely many conjugacy classes of D-subgroups of $MCG(S)$.*

PROOF. There are finitely many possible topological types of surfaces S' that can appear in a covering map $p : S \rightarrow S'$. Fix one such S' . Unless S' is a disk with 2 holes or a projective plane with 2 holes, choose two tight simple closed curves α and β that fill S' (this means that each complementary component is a disk or an annulus that contains a boundary component of S'). If S' is a disk with 2 holes choose an immersed curve α with one self-intersection point so that all 3 complementary components are annuli containing one boundary component of S' . Similarly, when S is a projective plane with 2 holes choose α so that it has two self-intersection points and the complementary components are annuli. We can view this as a kind of a cell structure on S' except that some of the cells have a hole in the interior.

If $p : S \rightarrow S'$ is a covering map we can pull back this “cell structure” to S . The cells are labeled by the corresponding cells in S' . There are only finitely many labeled cell complexes that arise in this fashion. Now suppose that $p_1, p_2 : S \rightarrow S'$ are two covering maps with the associated labeled cell complexes isomorphic via a homeomorphism that preserves labels. This yields $h : S \rightarrow S$ such that $p_1 = p_2h$, and this means that $D(p_1)$ and $D(p_2)$ are conjugate in $MCG(S)$: $D(p_1) = h^{-1}D(p_2)h$. \square

3. OUTLINE OF PROOF OF THEOREM 1.1

Suppose S is orientable. Since $MCG(S)$ is not virtually abelian, S is not S^2 minus at most three points. If S is S^2 minus four points or T^2 minus at most one point, then $MCG(S)$ is virtually free. In those cases, the curve complex X (see Appendix) is not connected, and we need to argue differently. One way is to modify the definition of X (see Appendix), then we can apply the same argument as follows. We omit the details.

The mapping class group $MCG(S)$ acts on the curve complex $X = X(S)$ which is hyperbolic by the celebrated theorem of Masur-Minsky [MM1]. For

finite oriented paths w, α in X write $|\alpha|_w$ for the maximal number of non-overlapping translates of w in α , and by $|w|$ denote the length of w . If W is an integer with $0 < W < |w|$, for any two vertices $x, y \in X$ define

$$c_{w,W}(x, y) = d(x, y) - \inf_{\alpha} (|\alpha| - W|\alpha|_w)$$

with α ranging over all paths from x to y . Finally, define

$$h_w : MCG(S) \rightarrow \mathbb{R}$$

by

$$h_w(g) = c_{w,W}(x_0, g(x_0)) - c_{w^{-1},W}(x_0, g(x_0))$$

where $x_0 \in X$ is a fixed base vertex.

Then h_w is a quasi-homomorphism [Fu, Proposition 3.10]. By construction, h_w is bounded on the stabilizer of every vertex x of X (i.e. the stabilizer of a scc in $MCG(S)$). Indeed, $0 \leq c_{w,W}(x_0, g(x_0)) \leq d(x_0, g(x_0)) \leq 2d(x_0, x)$, so $|h_w(g)| \leq 2d(x_0, x)$ for any $g \in MCG(S)$ that fixes x .

In [BeFu, Proposition 11] we showed that the action of $MCG(S)$ on X satisfies a certain technical condition called WPD (see Appendix). In the presence of this condition, the following two statements about hyperbolic elements g_1, g_2 are equivalent [BeFu, Proposition 6]:

- some positive powers $g_1^{n_1}$ and $g_2^{n_2}$ are conjugate, and
- for any quasi-axes ℓ_1 and ℓ_2 of g_1 and g_2 respectively there is a constant $C > 0$ such that for any $L > 0$ there are segments $\Sigma_i \subset \ell_i$ of length L so that a translate of Σ_1 is contained in the C -neighborhood of Σ_2 and the orientations of the two segments are parallel.

Write $g_1 \sim g_2$ if the two statements hold. We then constructed in [BeFu, Proposition 2] an infinite sequence $f_1, f_2, \dots \in MCG(S)$ of hyperbolic elements such that

- $f_i \not\sim f_i^{-1}$ for all i , and
- $f_i \not\sim f_j^{\pm 1}$ for $i \neq j$.

If $C = \langle \phi \rangle$ is a cyclic subgroup of $MCG(S)$ generated by a hyperbolic element ϕ then $\phi \sim f_i^{\pm 1}$ for at most one i . So if we are given a finite collection of cyclic subgroups generated by hyperbolic elements we may assume (after removing a finite subset of the f_i 's) that they do not contain any conjugates of any nontrivial powers of the f_i 's. Now, as in the proof of [BeFu, Theorem 1], inductively choose paths w_1, w_2, \dots where w_i is a long segment along a quasi-axis of f_i so that $h_{w_i,W} : MCG(S) \rightarrow \mathbb{R}$ is unbounded on $\langle f_i \rangle$ but bounded on the given cyclic subgroups and on $\langle f_1 \rangle, \dots, \langle f_{i-1} \rangle$ (when the cyclic subgroup is generated by an element that is not hyperbolic, then any $h_{w,W}$ is bounded on it). Then the quasi-homomorphisms $h_{w_i,W}$ give an infinite linearly independent collection in $QH(MCG(S))$ and by construction they are all bounded on the subgroups in (1) and (2).

Suppose S is non-orientable. Since $MCG(S)$ is not virtually abelian, S is not \mathbb{RP}^2 minus at most two points (MCG is finite, [Ko]), nor Klein bottle minus at most one point (MCG is finite (no puncture) or virtually \mathbb{Z} (one puncture) [Stu]). As we show in the appendix, the curve complex X is connected, and delta-hyperbolic, and the action of MCG is WPD. Therefore the same argument applies.

4. EXTENSION OF THEOREM 1.1

LEMMA 4.1. *Let $p : S \rightarrow S'$ be a covering map between compact connected surfaces of negative Euler characteristic, possibly with boundary, possibly nonorientable. Then the image of*

$$p^* : X(S') \rightarrow X(S)$$

is a quasi-convex subset of $X(S)$. Namely, there exists a constant P such that a geodesic (in $X(S)$) between any two points in $p^(X(S'))$ is in the P -neighborhood of $p^*(X(S'))$.*

This lemma complements the result of Rafi and Schleimer [RS] that p^* is a quasi-isometric embedding. In the case when S' is orientable the statement also follows from [RS].

PROOF. We use terminology and results from Appendix. Let a, b be two vertices of $X(S')$ and let c be a scc as in Lemma 5.7. Likewise, let \tilde{a}, \tilde{b} be the preimages of a, b in S , viewed as multi-curves. Let \tilde{c} be the preimage of c . Since the length $l_{ab}(c)$ of c is bounded above by $R\sqrt{I(a, b)}$ it follows that

$$l_{\tilde{a}\tilde{b}}(\tilde{c}) \leq |p|R\sqrt{I(a, b)} = \sqrt{|p|}R\sqrt{I(\tilde{a}, \tilde{b})}$$

and similarly

$$\begin{aligned} m_{\tilde{a}\tilde{b}}(\tilde{c}) &= \sup \frac{I(\tilde{c}, \tilde{x})}{l_{\tilde{a}\tilde{b}}(\tilde{x})} = \sup \frac{I(c, x)}{l_{ab}(x)} = m'_{ab}(c) \\ &\leq R/\sqrt{I(a, b)} = \sqrt{|p|}R/\sqrt{I(\tilde{a}, \tilde{b})} \end{aligned}$$

where \tilde{x} runs over scc's in S and x denotes $p(\tilde{x})$ (the modulus can be computed by running the sup over all immersed curves, see Remark 5.8). Therefore, \tilde{c} is within a bounded distance from $\text{Mid}(\tilde{a}, \tilde{b})$ (see Remark 5.9). A similar argument applies to $\text{Mid}(\tilde{a}, \tilde{b}; \frac{p}{q})$. \square

In the orientable case the following proposition is a consequence of [MS, Theorem 2].

PROPOSITION 4.2. *Let S be a compact surface that admits pseudo-Anosov homeomorphisms. Then there is a pseudo-Anosov homeomorphism $\phi : S \rightarrow S$ such that no nontrivial power of ϕ (is conjugate within $MCG_{\pm}(S)$ to a homeomorphism that) descends down a proper covering $S \rightarrow S'$.*

PROOF. In this proof it is convenient to collapse boundary components to punctures. If S is a 4 times punctured sphere, a once punctured torus, or a once punctured $\#_1^3 \mathbb{R}P^2$ then S admits a pseudo-Anosov homeomorphism, and such a homeomorphism cannot descend down a proper covering map since simpler surfaces do not admit such homeomorphisms. (For the first two cases $\chi = -1$ so the assertion is obvious. For the last case this takes a thought since $\chi = -2$, so conceivably there is a double cover. But the covered surface would have to have one puncture and under a double cover a single puncture must lift to two punctures for homological reasons.)

In all other cases we will construct $\phi : S \rightarrow S$ such that the stable and unstable foliations of ϕ have a unique singularity with > 2 prongs. This will imply that ϕ and its powers do not descend down a proper covering map. Note that if we have such a map ϕ for a surface S with p punctures, then we also have it for the surface S with $> p$ punctures, by taking an appropriate power of ϕ and declaring fixed points to be punctures. Thus it suffices to construct ϕ in the “minimal” cases.

Recall Thurston’s construction of pseudo-Anosov homeomorphisms of a surface S [Th]. Let α and β be two tight collections of pairwise disjoint essential 2-sided simple closed curves in S such that $\alpha \cup \beta$ fills S . By composing Dehn twists in these curves one obtains a pseudo-Anosov homeomorphism $\phi : S \rightarrow S$ whose (un)stable foliation has one singularity for every complementary component of $\alpha \cup \beta$. If the component is a $2k$ -gon, then the singularity is k -pronged (so if $k = 2$ it’s not a singularity, and if $k = 1$ there must be a puncture in this component). The proof is now contained in Figures 1-8.

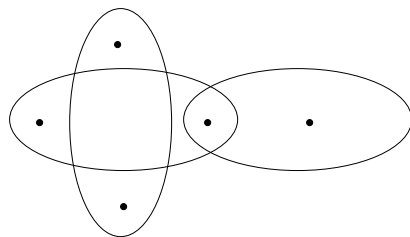


FIGURE 1. S^2 with 5 punctures. Each puncture is a 1-pronged singularity, and there is a 3-pronged singularity.

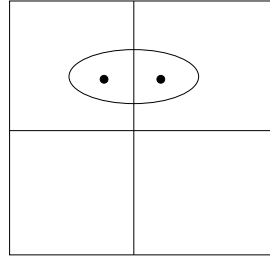


FIGURE 2. T^2 with 2 punctures, pictured as the square with opposite sides identified. The curves are the meridian and the longitude as well as a curve surrounding the two punctures. Each puncture is a 1-pronged singularity, and there is a 4-pronged singularity.

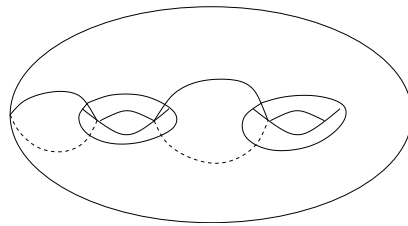


FIGURE 3. Genus 2 surface with a single 6-pronged singularity. In a similar way we get a genus $g \geq 2$ surface with a single $4g - 2$ -pronged singularity.

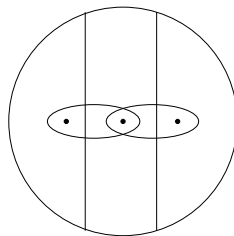


FIGURE 4. $\mathbb{R}P^2$ with 3 punctures, pictured as a disk with antipodal points on the boundary identified. Each puncture is a 1-pronged singularity, and there is a 3-pronged singularity.

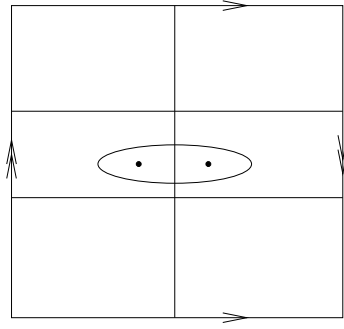


FIGURE 5. The Klein bottle K with 2 punctures. Each puncture is a 1-pronged singularity, and there is a 4-pronged singularity.

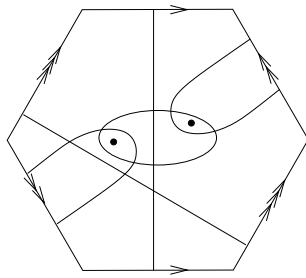


FIGURE 6. Surface with $\chi = -1$ and 2 punctures. Each puncture is a 1-pronged singularity, and there is a 6-pronged singularity.

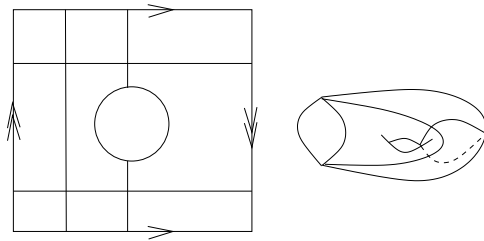


FIGURE 7. $K \# T^2$ with a single 4-pronged singularity. In a similar way we get $K \# \#_1^n T^2$ with a single $4n$ -pronged singularity.

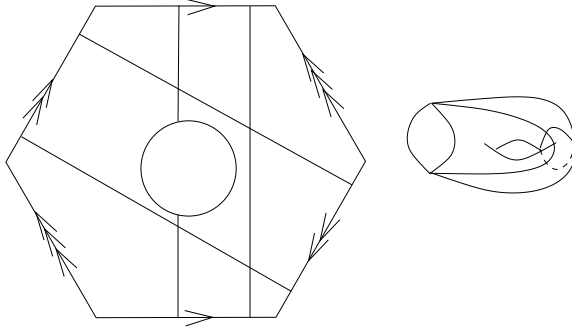


FIGURE 8. $\#_1^3 \mathbb{R}P^2 \# T^2$ with a single 6-pronged singularity. In a similar way we get $\#_1^3 \mathbb{R}P^2 \# \#_1^n T^2$ with a single $4n + 2$ -pronged singularity. \square

COROLLARY 4.3. *Let $\phi : S \rightarrow S$ be as in Proposition 4.2 and let g belong to the centralizer $Z(\phi)$ of $\langle \phi \rangle$ in $MCG(S)$. If g has finite order and is realized as an isometry of a hyperbolic structure on S then this isometry necessarily has a fixed point.*

PROOF. Fix a hyperbolic structure on S such that g is realized by an isometry $S \rightarrow S$. Let Λ be the stable geodesic lamination associated with ϕ . Then $g(\Lambda)$ is the stable geodesic lamination associated with $g\phi g^{-1} = \phi$ and so $g(\Lambda) = \Lambda$. By construction, there is $k > 2$ and a unique complementary component of Λ which is a k -gon. Therefore g maps this component to itself and has a fixed point. \square

We will need a small generalization of this.

COROLLARY 4.4. *Suppose $\phi : S \rightarrow S$ is a pseudo-Anosov homeomorphism whose stable foliation has precisely n k -pronged singularities for a certain $k > 2$. Let Δ be a finite subgroup of the centralizer $Z(\phi)$, and suppose that Δ is realized as a group of isometries of a hyperbolic metric on S . If $|\Delta| > n$ then Δ does not act freely on S .*

PROOF. Now Δ permutes the complementary k -gons and $|\Delta| > n$ ensures that the stabilizer of a k -gon is nontrivial, and fixes a point. \square

COROLLARY 4.5. *Fix a homeomorphism $\phi : S \rightarrow S$ as in Proposition 4.2. Let ℓ be a quasi-axis of ϕ in the curve complex $X(S)$. Then for every proper covering map $p : S \rightarrow S'$ and every $B > 0$ there is $L > 0$ such that no translate of ℓ has a segment of length L that is contained in the B -neighborhood of $p^*(X(S'))$.*

PROOF. First suppose the covering map $p' : S \rightarrow S'$ is regular. Let Δ be the deck group. Then Δ acts simplicially on the curve complex $X(S)$ and the fixed set is precisely $p^*(X(S'))$ (it is a subcomplex of the barycentric subdivision of $X(S)$). Now suppose that some translate $h(\ell)$ has a segment of length L contained in the B -neighborhood of $p^*(X(S'))$. Then for every $g \in \Delta$ we have that $gh(\ell)$ and $h(\ell)$ have segments of length L that are within $2B$ of each other. From [BeFu] we conclude that when L is sufficiently large there is $N > 0$ such that $gh\phi^N h^{-1}g^{-1} = h\phi^N h^{-1}$. This means that g is in the centralizer of $h\phi^N h^{-1}$. But then Corollary 4.3 (applied to $h\phi^N h^{-1}$ and g) implies that when g acts as an isometry of a hyperbolic structure on S then it fixes a point. But by construction, g acts without fixed points.

Now consider the general case, when $p : S \rightarrow S'$ is not regular. Let $q : \tilde{S} \rightarrow S$ be a finite cover with \tilde{S} connected so that pq (and hence q) is regular. Let $\tilde{\Delta}$ be the deck group of pq and $\Delta \subset \tilde{\Delta}$ the deck group of q . Again suppose that $h(\ell)$ contains a segment of length L that is inside the B -neighborhood of $p^*(X(S'))$, for some $h \in MCG(S)$. For some $M > 0$ the homeomorphism $h\phi^M h^{-1}$ lifts to \tilde{S} . Fix one such lift $\tilde{\tau} : \tilde{S} \rightarrow \tilde{S}$ and note that

- the stable foliation of τ has precisely $|q|$ (= number of sheets) singularities with k prongs, for a certain $k > 2$, and
- $q^*(\ell) \subset X(\tilde{S})$ is a quasi-axis of τ .

Of course, a segment of length L in $q^*(\ell) \subset X(\tilde{S})$ is also contained in the B -neighborhood of $q^*(p^*(X(S')))$. It now follows exactly as above that for some $N > 0$ the deck group $\tilde{\Delta}$ is contained in the centralizer of $\tilde{\tau}^N$. Corollary 4.4 then says that $\tilde{\Delta}$ cannot act freely as an isometry group on \tilde{S} , contradiction. \square

THEOREM 4.6. *Let S be a compact surface such that $MCG(S)$ is not virtually abelian. Then $QH(MCG(S))$ is infinite-dimensional. Moreover, for every finite collection of cyclic subgroups $C_1, \dots, C_k \subset MCG(S)$ there is an infinite-dimensional subspace of $QH(MCG(S))$ such that each representative quasi-homomorphism h of any element in this subspace satisfies:*

- (1) h is bounded on each C_i ,
- (2) h is bounded on the stabilizer $H(\alpha)$ of every essential (non-degenerate when S is non-orientable) simple closed curve α in S , and
- (3) h is bounded on every D -subgroup $D(p)$.

PROOF. The strategy of proof is the same as that of Theorem 1.1 but certain additional things need to be arranged.

Let $\phi : S \rightarrow S$ be the pseudo-Anosov homeomorphism from Proposition 4.2. Also choose another pseudo-Anosov homeomorphism $\psi : S \rightarrow S$ so that ϕ, ψ do not have isotopic nontrivial powers. The subgroup $F \subset MCG(S)$ generated by high powers ϕ^N, ψ^N of ϕ and ψ is a nonabelian free group and

moreover an F -equivariant map $j : T \rightarrow X$ is a quasi-isometric embedding, where T is the Cayle tree of F with respect to ϕ^N, ψ^N . Thus for every $1 \neq f \in F$ we obtain a quasi-axis for f by taking the axis of f in T and mapping it to X by j . There are constants K_0, C_0 so that in this way we obtain a (K_0, C_0) -quasi-axis regardless of which $1 \neq f \in F$ we chose. All paths w used in the construction of a quasi-homomorphism $h_{w,W}$ will be j -images of edge-paths in T , and in particular they are all (K_1, C_1) -quasi-geodesic for a fixed K_1, C_1 . It now follows that every minimizer path is a (K, C) -quasi-geodesic for a fixed K, C (see the proof of Theorem 1.1 given in Section 3).

Let D_1, \dots, D_m be representatives of conjugacy classes of D -subgroups of $MCG(S)$ (see Lemma 2.1). If a quasi-homomorphism is bounded on each D_k then it is bounded on every D -subgroup. Say $D_k = D(p_k : S \rightarrow S'_k)$ and let $A_k = p_k^*(X(S'_k)) \subset X(S)$. Then A_k is quasi-convex in $X(S)$ (see Lemma 4.1). Choose $B > 0$ such that any (K, C) -quasi-geodesic joining two points of A_k is contained in the B -neighborhood of A_k , $k = 1, \dots, m$ and choose $L > 0$ such that no segment in ℓ (the quasi-axis of ϕ) can be translated into the B -neighborhood of any A_k (see Corollary 4.5).

Now choose $f_1, f_2, \dots \in F$ so that in addition to $f_i \not\sim f_j^{\pm 1}$ for $i \neq j$ and $f_i \not\sim f_j^{-1}$, the quasi-axis of each f_i contains a path in a translate of ℓ of length $> L$. Finally, choose a path w_i in the quasi-axis of f_i so that in addition to the previous requirements w_i contains a segment of length $> L$ which is a translate of a segment in ℓ . It follows that no w_i can be translated into the B -neighborhood of any A_k . In particular, (K, C) -quasi-geodesics joining points of A_k do not contain copies of w_i . If the base vertex x_0 belongs to A_k then we deduce that $h_{w_i,W}(g) = 0$ for every $g \in MCG(S)$ that leaves A_k invariant. When x_0 is moved, $h_{w,W}$ changes by a bounded amount. \square

COROLLARY 4.7. *Suppose $MCG(S)$ is not virtually abelian. Let Q_1, \dots, Q_m be a finite collection of subgroups of $MCG(S)$ such that each Q_i is either cyclic, or the stabilizer of a non-degenerate scc, or a D -subgroup. Then*

$$Q_1 Q_2 \cdots Q_m \neq MCG(S).$$

Indeed, any finite union of sets of the above form is a proper subset of $MCG(S)$.

5. APPENDIX

In this Appendix we discuss the curve complex X of a compact surface S , which is orientable or non-orientable. Masur and Minsky [MM1] have shown that X is Gromov-hyperbolic when S is orientable. Bowditch [Bo1] has given another argument for that fact using $CAT(0)$ geometry. We follow his argument and explain that it also applies to non-orientable surfaces with minor changes.

The mapping class group $MCG(S)$ acts on X by isometries. When S is orientable, Masur and Minsky have shown that an element in $MCG(S)$ acts by a hyperbolic isometry on X if and only if it is pseudo-Anosov. It is shown that the action is weakly properly discontinuous [BeFu]. We explain that those facts remain true when S is non-orientable. The argument does not require any change.

5.1. *The curve complex.* Let S be a closed surface equipped with n punctures (i.e. distinguished points).

When S is orientable, let $X = X(S, n)$ be the simplicial complex whose vertices are isotopy classes of *essential* scc's (i.e. those that don't bound a disk or a punctured disk).

When S is non-orientable, vertices are isotopy classes of scc's *non-degenerate* (does not bound a disk, a punctured disk, or a Möbius band). If c bounds a Möbius band, the core curve of the Möbius band is non-degenerate.

In either case, a collection of vertices of X bounds a simplex provided representative curves can be found that are pairwise disjoint.

Denote by $I(a, b)$ the minimal intersection number between representatives. It is attained when there are no bigons in the complement.

When $S = S^2$ and $n \leq 3$ then $X = \emptyset$. When $S = S^2$ and $n = 4$, or $S = T^2$ and $n \leq 1$, then X is a discrete set (but see Remark 5.4 below). So we will assume that $n \geq 5$ when $S = S^2$ and $n \geq 2$ when $S = T^2$. Similarly, we will assume $n \geq 3$ when S is $\mathbb{R}P^2$, and we assume $n \geq 2$ when S is the Klein bottle.

LEMMA 5.1. *X is connected.*

PROOF. Let a, b be two vertices. We will connect them by an edge-path.

If $I(a, b) = 0$ then there is an edge between them. Next, we consider the case that both a and b are 2-sided. If $I(a, b) = 1$ then the regular neighborhood N of $a \cup b$ is a torus with one boundary component c . This is an essential curve since otherwise we are in a torus with $n \leq 1$. If c bounds a Möbius band, replace c by the core curve. Thus c is connected to both a and b .

As the next case, suppose we can find two consecutive intersection points on a that have the same sign (orient a and transversely orient b , so signs make sense). Form a curve c as the union of the arc in a that joins these two points and has no intersection points in the interior, and of an arc in b joining these two points. Then either $I(b, c) = 1$ or $I(b, c) = 0$ and c is one-sided. In either case, c is non-degenerate. Since $I(a, c) < I(a, b)$ the claim now follows by induction.

Now suppose that we have two intersection points of the same sign on a with one other intersection point in between. Then we can choose an arc on b connecting the two points that does not intersect the interior of the arc on a , so the same argument works.

Finally, if a and b intersect in two points of opposite signs, then the regular neighborhood N is a sphere with 4 boundary components. At least one boundary component must be essential, and if it is degenerate replace it by the core of the Möbius band it bounds, so $\text{dist}(a, b) = 2$.

Now suppose at least one of a, b , say b , is 1-sided. If $I(a, b) = 1$ then the regular neighborhood N of $a \cup b$ is either $\mathbb{R}P^2$ with two boundary components or the Klein bottle with one boundary component. In either case, at least one boundary component c must be essential. If c is degenerate, replace it by the core of the Möbius band it bounds. Thus $d(a, b) = 2$.

Now suppose $I(a, b) > 1$. Fix two consecutive intersection points on a . Using the arc in a between the two points (and contains no other intersection points) and an arc in b between the two points (there are two such arcs in b), form a scc b_1 . Using the other arc in b , form a scc b_2 . By construction, $I(b_1, b), I(b_2, b) \leq 1$, and also $I(b_1, a), I(b_2, a) < I(b, a)$.

Note that $b_1 + b_2 = b$ in $H^1(S, \mathbb{Z}_2)$. A scc is one-sided if and only if its square is non-trivial in $H^2(S, \mathbb{Z}_2)$. By our assumption, b^2 is non-trivial in $H^2(S, \mathbb{Z}_2)$. Therefore, one of b_1, b_2 , say b_1 , is one-sided. It follows that b_1 is non-degenerate, which gives a vertex in X , so we are done by induction. \square

Higher connectivity of the complex of curves for orientable and non-orientable S is discussed in [Iv] (for example, Theorem 2.6).

REMARK 5.2. Induction shows easily that $\text{dist}(a, b) \leq 2I(a, b)$. This argument is essentially contained in [L, Lemma 2]. Actually, the argument can be improved. In the generic case above we can choose one of two possible arcs in b connecting the two intersection points. We choose the one with fewer intersection points. This guarantees that $I(a, c) < I(a, b) - \lfloor \frac{I(a, b)-1}{2} \rfloor$. Induction then shows that $\text{dist}(a, b) \leq O(\log I(a, b))$.

REMARK 5.3. There is no lower bound to $\text{dist}(a, b)$ in terms of $I(a, b)$. Just think about two curves that intersect a lot but don't fill. In fact, it is nontrivial (and involves Nielsen-Thurston theory) to show that X has infinite diameter.

REMARK 5.4. When $S = T^2$ and $n = 0$ or 1 , it is natural to change the definition and connect a to b when $I(a, b) = 1$. The resulting graph is classically known as the Farey graph – it is the 1-skeleton of the standard tessellation of \mathbb{H}^2 by ideal triangles, including the vertices at infinity. There is classical number theory (continued fractions) associated with this graph – see the article by C. Series [Ser]. This graph is also hyperbolic. Exercise: Give a direct proof of this fact.

5.2. *Hyperbolicity of the curve complex.* Masur-Minsky [MM1] showed that the curve complex of an orientable surface is delta-hyperbolic. Bowditch [Bo1] gave another proof of this fact. We outline his argument and explain that his argument applies to non-orientable surfaces as well.

By a *curve system* we mean a collection of essential, pairwise disjoint scc's (we allow parallel curves).

We will think of S as a compact surface with punctures. Let a, b be two curve systems on the surface S . Realize them so that they intersect minimally (every bigon contains a puncture). We will assume that $a \cup b$ is *filling*, i.e. each disk in the complement contains a puncture, and we will construct a particular Euclidean metric on S with cone singularities. All cone angles will be multiples of π and if an angle equals π then the point is one of the punctures. Realize a and b in S so that they intersect minimally (every bigon contains a puncture). Take a collection of squares, one for each intersection point, and glue them in the obvious way so that $a \cup b$ is the dual 1-skeleton of the resulting 2-complex. This complex is S . Also, the horizontal and vertical foliations on the squares match up and give a pair of (measured) foliations on S . The area of S equals the intersection number $I(a, b)$. By $l_S(c)$ denote the length (with respect to this metric) of a shortest curve representing c .

Now set:

- the *length* of a curve c is $l(c) = l_{ab}(c) = I(a, c) + I(b, c)$,
- the *modulus* of c is $m(c) = m_{ab}(c) = \sup\{\frac{I(c, x)}{l(x)} \mid x \text{ is a curve}\}$.

We will restrict ourselves to the case when all curves in a and b are 2-sided. The reason is that we want to be able to talk e.g. about the curve system Na for $N = 1, 2, 3, \dots$, and this can be taken to mean “take N parallel copies of a ”. If a were 1-sided, we could only do this formally.

LEMMA 5.5. *Suppose each curve in a and b is 2-sided. Then*

$$\frac{1}{\sqrt{2}}l_S(c) \leq l_{ab}(c) \leq 2l_S(c).$$

PROOF. First suppose that the geodesic representing c is transverse to a and b . The first inequality follows from the fact that each complementary component of $a \cup b$ has diameter $\sqrt{2}$, and the second from the fact that the distance between two components of a (and of b) is ≥ 1 , so that $I(a, c) \leq l_S(c)$ and $I(b, c) \leq l_S(c)$.

Now suppose that c is one of the components of a (say). Then $l_S(c) = I(b, c) = I(a, c) + I(b, c) = l_{ab}(c)$. \square

REMARK 5.6. If a is 1-sided curve and $c = a$ then $l_S(c) = I(b, c)$ while $l_{ab}(c) = I(a, c) + I(b, c) = 1 + I(b, c)$, so the lemma would still hold; however, it would fail if we allow formal scaling for $c = Na$.

LEMMA 5.7. *Assume that if S is S^2 then $n \geq 4$, if $S = \mathbb{R}P^2$ then $n \geq 3$ and if S is the Klein bottle then $n \geq 2$. There is a constant $R = R(S, n)$ such that for any filling pair a, b of 2-sided curve systems there is a curve c with length bounded above by $R\sqrt{I(a, b)}$ and modulus bounded above by $R/\sqrt{I(a, b)}$. The curve c is 2-sided, essential, possibly degenerate.*

PROOF. (Outline) We follow Bowditch [Bo1]. The argument requires only minor changes when S is non-orientable.

STEP 1. We will find an essential scc c with length $l_S(c)$ bounded by $C\sqrt{\text{Area}(S)}$ and with an annular (possibly Möbius band when c is one-sided) neighborhood around it of width at least $\sqrt{\text{Area}(S)}/C$ on each side (C depends only on S, n). This will imply that for each x we have $l_S(x) \geq I(x, c)\sqrt{\text{Area}(S)}/C$ and this implies that $m_{ab}(c) \leq R/\sqrt{I(a, b)}$. Thus if c is 2-sided we are done, and if c is 1-sided replace it by the boundary of a regular neighborhood.

STEP 2. Quadratic isoperimetric inequality holds in S . This means that if D is a disk in S that contains at most one puncture, then $\text{Area}(D) \leq C \text{length}(\partial D)^2$ where C can be taken to be twice the constant for euclidean plane. Indeed, when D has no punctures this follows from the CAT(0) theory, and when there is one puncture, pass to the double cover. Note that the isoperimetric inequality fails when there are two punctures in D (think of the double of an infinite half-strip).

STEP 3. A *spine* in S is a graph $\Gamma \subset S \setminus \{\text{punctures}\}$ such that inclusion is π_1 -surjective. We claim that there is a number $\eta_1 > 0$ (as a function of the number of punctures) such that the length of any spine satisfies the isoperimetric inequality (II)

$$\text{length}(\Gamma) \geq \eta_1 \sqrt{\text{Area}(S)}.$$

It is enough to prove this for “minimal spines”, i.e. those that are no longer spines after an edge is removed. When S has no punctures, a minimal spine has one complementary component, which is a disk, and the standard II says:

$$\text{Area}(S) \leq C \text{length}(\partial D)^2 = \frac{C}{4} \text{length}(\Gamma)^2$$

which proves the claim. When there are n punctures, the complement consists of n disks D_i containing one puncture each. From II we see that

$$\text{Area}(S) = \sum \text{Area}(D_i) \leq C \sum \text{length}(\partial D_i)^2$$

and so at least one D_i must have long boundary, implying that Γ is long.

STEP 4. Let N be the length of a longest chain of connected incompressible, pairwise non-isotopic subsurfaces of $S \setminus \{\text{punctures}\}$ and set $\eta_0 = \eta_1/(100 + 2N)$. We claim that there are two non-degenerate simple closed curves x, y in $S \setminus \{\text{punctures}\}$ whose distance is $\geq \eta_0 \sqrt{\text{Area}(S)}$.

For a curve a denote by $\text{length}'(a)$ the length of a when a is 2-sided, and $\frac{3}{2} \text{length}(a)$ when a is 1-sided. Let x be a non-degenerate scc in $S \setminus \{\text{punctures}\}$ with $\text{length}'(x)$ within η_0 of $\inf\{\text{length}'(a) \mid a \text{ is non-degenerate}\}$. We will show that there is a non-degenerate scc y outside the $\eta_0 \sqrt{\text{Area}(S)}$ -neighborhood of x . Suppose not and consider the growing family of neighborhoods N_t , $0 \leq t \leq \eta_0 \sqrt{\text{Area}(S)}$ of x . At $\leq N$ times the topology of the

minimal incompressible subsurface that contains N_t changes. Each critical point can be connected to x by two arcs of length $\leq \eta_0 \sqrt{\text{Area}(S)}$ so that the obtained 1-complex Γ consisting of these (disjoint, except for endpoints) arcs and of x carries π_1 . The last surface is a spine by our assumption, so this 1-complex is a spine. Thus,

$$\text{length}(x) + 2N\eta_0 \sqrt{\text{Area}(S)} \geq \eta_1 \sqrt{\text{Area}(S)}$$

and we deduce

$$\text{length}(x) \geq 100\eta_0 \sqrt{\text{Area}(S)}.$$

Now it's not hard to find a curve in Γ much shorter than x . More explicitly, we have the following cases:

- x is 1-sided. If Γ is a circle, then S is $\mathbb{R}P^2$ with ≤ 1 puncture. If there is a short arc in Γ whose endpoints are far apart along x , then this arc together with one of the two arcs of x forms a scc x' with $I(x, x') = 1$ which is much shorter than x . In particular, x' is non-degenerate and 1-sided. If Γ is the union of x and only one short arc, then S is either $\mathbb{R}P^2$ with ≤ 2 punctures or the Klein bottle with ≤ 1 puncture. Now suppose that there are ≥ 2 short arcs in Γ and each has endpoints close to each other along x . For each such short arc consider the scc obtained by adding the short subarc of x . If one such scc is nondegenerate then we have a contradiction to the choice of x . Otherwise each such scc bounds a punctured disk or a Möbius band. In the latter case, the core would be a better choice for x . If two of these curves bound punctured disks then their "connect sum" along an arc of x produces a curve x' which bounds a twice punctured disk and whose length is at most slightly bigger than that of x . Thus $\text{length}'(x') \leq \frac{2}{3} \text{length}'(x) + \epsilon$ and x' is a better choice than x .
- x is 2-sided and nonseparating. Then there is a short arc in Γ that joins x to itself from opposite sides. Let x' be the curve formed by this arc together with the shorter of the two subarcs of x . Then x' is much shorter than x (even $\frac{3}{2} \text{length}(x')$ is much shorter than x in case x' is 1-sided) and is non-degenerate because $I(x, x') = 1$.
- Now suppose x is separating. In particular, x is 2-sided. If there is a short arc γ in Γ such that $x \cup \gamma$ has a nonorientable regular neighborhood, then γ union the shorter of the two subarcs in x contradicts the choice of x . Similarly, if there is γ whose endpoints are far apart on x then both curves formed by γ and an arc in x have about half the length of x , so they have to be degenerate. If one bounds a Möbius band then the core would be a better choice for x , so necessarily in this case each bounds a punctured disk, so the component of $S - x$ containing $\gamma - x$ is necessarily a twice punctured disk. Next, if Γ contains two short arcs γ_1 and γ_2 with each pair of endpoints close along x , and the endpoints of γ_1 are not nearly antipodal from the endpoints

of γ_2 , then we find x' that bounds a twice punctured disk as in the first bullet.

There are now two cases that remain.

- Γ has one short arc on one side of x (which is a twice punctured disk) and the endpoints are far apart, and on the other side it has two short arcs, with endpoints close, but nearly antipodal from each other. Now use the arc from the twice punctured disk side and an arc on the other side to construct x' which is a better choice than x .
- Both sides of x are twice punctured disks (and so S is a 4-times punctured sphere). If necessary, homotope x to a geodesic. Now each side is a twice punctured disk with totally geodesic boundary. Double branched cover, branched over the punctures, is a flat annulus, and the punctured disk is the quotient of the obvious involution with two fixed points. The short arc of Γ is, up to isotopy, any of the geodesic arcs perpendicular to both boundary components. There is a 1-parameter family of such arcs and we may arrange that on one side the endpoints are close, and on the other side at least one of the two endpoints is close to the first two. Now we may construct a scc whose length is only slightly longer than $\frac{1}{2} \text{length}(x)$ and bounds a twice punctured disk.

STEP 5. Let M be the maximal number of non-degenerate, pairwise disjoint, nonparallel scc's in $S \setminus \{\text{punctures}\}$. We claim that there is an essential annulus whose boundary components are at least $\frac{\eta_0}{M} \sqrt{\text{Area}(S)}$ apart. To see this, consider a 1-Lipschitz map $S \rightarrow [0, d(x, y)]$ that sends x to 0 and y to $d(x, y)$. Subdivide $[0, d(x, y)]$ into M subintervals and consider the $M + 1$ preimages of the vertices. Without loss the map is transverse to the vertices. First note that each preimage must contain a component which is non-degenerate (otherwise x and y would be separated by a union of degenerate curves). By the choice of M , among them two are parallel.

STEP 6. (Besicovich Lemma) Let A be an annulus and $f : A \rightarrow [0, d]$ a 1-Lipschitz map. Then there is a core curve in A of length $\leq \text{Area}(A)/d$. Imagining that f is smooth, this can be seen by defining a 2-form ω on A that evaluates on (v, w) where $df(v) = 0$ as $\pm |v| df(w)$ (the sign is determined by orientations). Then the 1-Lipschitz condition implies that ω is a subarea form, and hence

$$\text{Area}(A) \geq \int_A \omega = \int_0^d \text{length}(f^{-1}(t)) dt$$

which means that for some t , $\text{length}(f^{-1}(t)) \leq \text{Area}(A)/d$.

Put $c = f^{-1}(t)$. Then c is an essential scc with length $\leq M \sqrt{\text{Area}(S)}/\eta_0$ and width $\geq \eta_0 \sqrt{\text{Area}(S)}/M$. \square

REMARK 5.8. We could define modulus of a curve c by taking the sup of $\frac{I(c, x)}{l(x)}$ as x runs over all *immersed* curves in the surface. Call the resulting

number $m'(c)$. Then *a priori* $m(c) \leq m'(c)$, but the proof of Lemma 5.7 shows that $m'(c) \leq R/\sqrt{I(a,b)}$.

In what follows it is best to think of 1-sided curves as being interchangeable with the associated degenerate curves. Sometimes one is more suitable than the other. To formalize this, consider the complex X' whose vertices are isotopy classes of essential 2-sided scc's and simplices correspond to collections of curves that can be realized disjointly. Then $X \cong X'$ by the map which is “identity” except it sends 1-sided curves to the associated degenerate 2-sided curves.

REMARK 5.9. Note that if c, d are 2-sided, $m(c) \leq R/\sqrt{I(a,b)}$ and $l(d) \leq R\sqrt{I(a,b)}$ then c and d are close in X' . Indeed, $I(c,d) \leq l(d)m(c) \leq R^2$, so $\text{dist}(c,d) \leq 1 + I(c,d) \leq 1 + R^2$. The same holds if say $l(d) \leq 100R\sqrt{I(a,b)}$.

The *coarse midpoint* for $\{a,b\}$ is the set $\text{Mid}(a,b) \subset X'$ of all 2-sided essential curves c with $m_{ab}(c) \leq R/\sqrt{I(a,b)}$. In view of Remark 5.9, we could also work with the set $\text{Mid}'(a,b)$ of all curves c with $l_{ab}(c) \leq R\sqrt{I(a,b)}$. The two sets intersect and any point in Mid is within $1 + R^2$ of any point in Mid' . Thus both sets have diameter at most $2 + 2R^2$.

Recall that in the above discussion a and b are curve systems. For example, note that $\text{Mid}(Na, Nb) = \text{Mid}(a, b)$. We now define the *weighted midpoint* $\text{Mid}(a, b; \frac{p}{q})$ as $\text{Mid}(qa, pb)$. Thus $\text{Mid}(a, b) = \text{Mid}(a, b; 1)$. Finally, we define the *coarse geodesic* between a and b as

$$\Lambda_{ab} = \bigcup \left\{ \text{Mid} \left(a, b; \frac{p}{q} \right) \mid 0 < \frac{p}{q} < \infty \right\}.$$

Note that for N large any component of a is in $\text{Mid}(a, b; \frac{1}{N})$ and any component of b is in $\text{Mid}(a, b; N)$.

LEMMA 5.10. *Suppose $I(b, c) = 0$ and $I(a, b) = I(a, c) = I$. Then $\text{Mid}(a, b)$ and $\text{Mid}(a, c)$ are close (here a, b, c are 2-sided curve systems).*

PROOF. Let x be a curve with $m_{ab}(x) \leq R/\sqrt{I}$ and $l_{ab}(x) \leq R\sqrt{I}$. Then

$$I(c, x) \leq m_{ab}(x)l_{ab}(c) \leq \frac{R}{\sqrt{I}}I = R\sqrt{I}$$

so $l_{ac}(x) \leq 2R\sqrt{I}$ and the claim follows from Remark 5.9. \square

Now suppose a, b, c are three curve systems, each pair filling. Consider the curve systems

$$\tilde{a} = I(b, c)a, \quad \tilde{b} = I(c, a)b, \quad \tilde{c} = I(a, b)c$$

and note that the intersection number between each pair is

$$\tilde{I} = I(a, b)I(a, c)I(b, c).$$

LEMMA 5.11. *The three sets $\text{Mid}(\tilde{a}, \tilde{b})$, $\text{Mid}(\tilde{b}, \tilde{c})$, $\text{Mid}(\tilde{c}, \tilde{a})$ are close to each other.*

PROOF. Let $x \in \text{Mid}(\tilde{a}, \tilde{b})$ be a curve from Lemma 5.7 (recall that if x bounds a Möbius band the core curve is in $\text{Mid}(\tilde{a}, \tilde{b})$). We estimate the length of x with respect to \tilde{a}, \tilde{c} . $I(x, \tilde{a}) \leq R\sqrt{\tilde{I}}$ by the choice of x , and the other estimate is as in Remark 5.9:

$$I(x, \tilde{c}) \leq m_{\tilde{a}\tilde{b}}(x)l_{\tilde{a}\tilde{b}}(\tilde{c}) \leq \frac{R}{\sqrt{\tilde{I}}}(I(\tilde{a}, \tilde{c}) + I(\tilde{b}, \tilde{c})) = 2R\sqrt{\tilde{I}}.$$

Thus

$$l_{\tilde{a}, \tilde{c}}(x) = I(x, \tilde{a}) + I(x, \tilde{c}) \leq 3R\sqrt{\tilde{I}}$$

and we are done by Remark 5.9. \square

By $\text{Center}(a, b, c)$ denote the union $\text{Mid}(\tilde{a}, \tilde{b}) \cup \text{Mid}(\tilde{b}, \tilde{c}) \cup \text{Mid}(\tilde{c}, \tilde{a})$. It is a nonempty set of uniformly bounded diameter.

LEMMA 5.12. *If $M \geq N$ then $\text{Mid}(M\tilde{a}, N\tilde{b})$ is close to $\text{Mid}(M\tilde{a}, N\tilde{c})$.*

PROOF. Let x be a special curve from Lemma 5.7 for the pair $M\tilde{a}, N\tilde{b}$. Then

$$I(x, N\tilde{c}) \leq m_{M\tilde{a}, N\tilde{b}}(x)l_{M\tilde{a}, N\tilde{b}}(N\tilde{c}) \leq \frac{R}{\sqrt{MN\tilde{I}}}(MN\tilde{I} + N^2\tilde{I}) \leq 2R\sqrt{MN\tilde{I}}$$

and we deduce as before that x is close to $\text{Mid}(M\tilde{a}, N\tilde{c})$. \square

The following lemma is motivated by Axiom (3) (see below). If $x \in \text{Mid}(a, b)$ we wish to show that $\text{Center}(a, b, x)$ and x are close. The issue is that x may intersect one of the curves, say a , more often than the other, so to compute $\text{Center}(a, b, x)$ we have to pass to multiples of a and b , with the coefficient of a being smaller than that of b . The point is that the effect on x is that it gets relatively shorter that way.

LEMMA 5.13. *Suppose $l_{ab}(x) \leq R\sqrt{I(a, b)}$, i.e. $x \in \text{Mid}'(a, b; 1)$. Then x is close to $\text{Mid}'(a, b; t)$ for all $t = \frac{p}{q}$ between 1 and $\frac{I(a, x)}{I(b, x)}$.*

PROOF. For concreteness assume

$$1 \leq \frac{p}{q} \leq \frac{I(a, x)}{I(b, x)}.$$

Then $l_{qa, pb}(x) = I(qa, x) + I(pb, x) = qI(a, x) + pI(b, x)$ and we wish to show

$$qI(a, x) + pI(b, x) \leq CR\sqrt{I(qa, pb)}$$

for some constant $C > 0$. Squaring and writing $I = I(a, b)$ we need to show

$$q^2I(a, x)^2 + p^2I(b, x)^2 + 2pqI(a, x)I(b, x) \leq C^2R^2pqI.$$

Since $l_{ab}(x) \leq R\sqrt{I}$ we certainly have $I(a, x) \leq R\sqrt{I}$ and $I(b, x) \leq R\sqrt{I}$, so the last term is bounded (with $C^2 = 2$ on the right hand side). The first term can be bounded using $q \leq p$ and $I(a, x) \leq R\sqrt{I}$, and the second using $pI(b, x) \leq qI(a, x)$ and then estimating as in the last term. \square

5.3. Hyperbolicity criterion. Let X be a connected graph with the edge-path metric. Suppose that for each pair a, b of vertices we have a subset Λ_{ab} with a “coarse order” (total, but not anti-symmetric), and $\Lambda_{ab} = \Lambda_{ba}$ as sets but with reversed order. For $x, y \in \Lambda_{ab}$ with $x \leq y$ write

$$\Lambda_{ab}[x, y] = \Lambda_{ab}[y, x] = \{z \in \Lambda_{ab} \mid x \leq z \leq y\}.$$

Also assume that $\phi : X^0 \times X^0 \times X^0 \rightarrow X^0$ is a ternary function on the vertices, symmetric under permutations, and such that $\phi(a, b, c) \in \Lambda_{ab}$. We assume that there is $K > 0$ so that

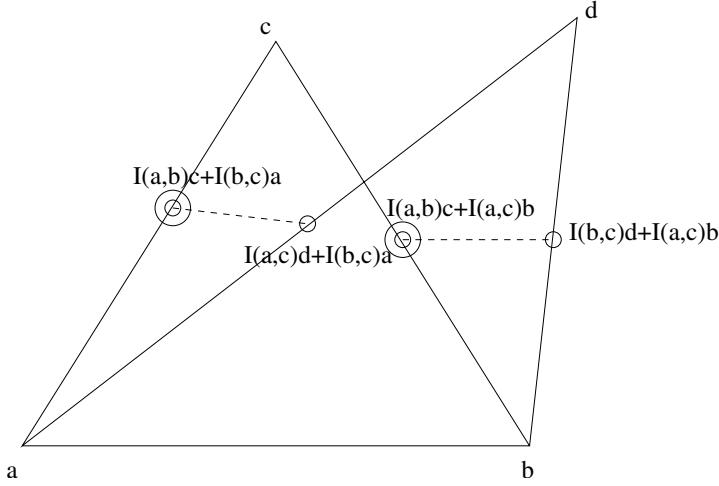
- (1) $\text{HausDist}(\Lambda_{ab}[a, \phi(a, b, c)], \Lambda_{ac}[a, \phi(a, b, c)]) \leq K$,
- (2) if c, d are adjacent vertices then $\text{diam}\Lambda_{ab}[\phi(a, b, c), \phi(a, b, d)] \leq K$,
- (3) if $c \in \Lambda_{ab}$ then $\Lambda_{ab}[c, \phi(a, b, c)] \leq K$.

THEOREM 5.14 ([Bo1, Proposition 3.1]). *It follows that X is Gromov hyperbolic. There exists Q such that for all a, b , Λ_{ab} is in the Q -neighborhood of a geodesic between a, b (we say Λ_{ab} is Q -quasi-convex).*

5.4. Verifications. We put $\phi = \text{Center}$. Axiom (1) follows from Lemma 5.12. Lemma 5.13 reduces the verification of Axiom (3) to the case $I(a, c) = I(b, c)$ when it follows immediately from definitions.

We now verify Axiom (2). For simplicity, we normalize so that $I(a, b) = I(a, d) = I(b, d) = I$ and we assume that $I(a, c) \geq I(b, c)$. Consider the point in $\text{Center}(a, b, c)$ that belongs to $\text{Mid}(I(a, b)c, I(b, c)a)$ (such a point is suggestively called $I(a, b)c + I(b, c)a$ in the figure below and it is represented by a “bull’s eye”). By Lemma 5.10 there is a point in $\Lambda[a, d]$ close to it, and a computation¹ shows that this point is $I(a, c)d + I(b, c)a$, using the same suggestive notation. Analogously, we can start with the point $I(a, b)c + I(a, c)b$ also representing $\text{Center}(a, b, c)$ and transfer it to $I(b, c)d + I(a, c)b$.

¹Note that Lemma 5.10 applies to qa, pb, pc , therefore $\text{Mid}(a, b; \frac{p}{q})$ is uniformly close to $\text{Mid}(a, c; \frac{p}{q})$ for any p, q . This gives a “linear” transfer map: $I(a, b)c + I(b, c)a = Ic + I(b, c)a = I(a, c)\frac{I}{I(a, c)}c + I(b, c)a$ which transfers to $I(a, c)d + I(b, c)a$.



The two bull's eyes are close (they are in $\text{Center}(a, b, c)$), and consequently the other two points are close. Note that one of them is “too high” and the other is “too low” (the center $\text{Center}(a, b, d)$ is represented by $a + d$ or $b + d$).

There is a 1-parameter family of points that interpolate between the two points above. By c_t denote the curve system $tc \cup (1-t)d$, so $c_1 = c$ and $c_0 = d$. Note that $I(c, c_t) = I(d, c_t) = 0$ for all t because c, d are 2-sided. Take the point $I(a, b)c + I(b, c)a$, transfer it to Λ_{a, c_t} , then to Λ_{b, c_t} , and finally to Λ_{bd} . When $t = 1$ we get $I(b, c)d + I(a, c)b$ and when $t = 0$ we get $I(a, c)d + I(b, c)b$, so for some $0 \leq t \leq 1$ we get the center for abd , proving the claim. The only thing we have to ensure is that the transfer from Λ_{a, c_t} to Λ_{b, c_t} is possible, i.e. that the point in question is on the “ c_t -side” of the centerpoint for abc_t . A little calculation² shows that the condition is that $I(a, c) \geq I(b, c)$.

5.5. *The action of $MCG(S)$ on X .* Let $MCG(S)$ be the mapping class group of S . Then $MCG(S)$ acts on X by isometries. Masur-Minsky ([MM1, Proposition 4.6]) showed a pseudo-Anosov element $g \in MCG(S)$ acts as a hyperbolic isometry on X in the following sense: there exists a quasi-geodesic γ in X which is invariant by g . In this case γ is called a (quasi-geodesic) axis of g . It is trivial that g is not hyperbolic if it is not pseudo-Anosov.

It is shown ([BeFu, Proposition 11]) that the action of $MCG(S)$ on X is weakly properly discontinuous (WPD) in the following sense: for every pseudo-Anosov element $g \in MCG(S)$, every $x \in X$, and every $C \geq 0$, there

²In a triangle xyz the point $\alpha x + \beta y = \frac{\alpha}{I(y, z)}I(y, z)x + \frac{\beta}{I(x, z)}I(x, z)y$ is on the x -side of the center iff $\frac{\alpha}{I(y, z)} \geq \frac{\beta}{I(x, z)}$ iff $\alpha I(x, z) \geq \beta I(y, z)$. Applying this to $c_t ab$ we get the condition $\frac{I \times I(a, c)}{tI(a, c) + (1-t)I}I(c_t, b) \geq I(b, c)I(a, b)$ i.e. $I \times I(a, c)(tI(b, c) + (1-t)I) \geq I(b, c)I(tI(a, c) + (1-t)I)$ which simplifies to $I(a, c) \geq I(b, c)$.

exists $N > 0$ such that the following set is finite:

$$\{h \in MCG(S) \mid \text{dist}(x, h(x)) \leq C, \text{dist}(g^N(x), hg^N(x)) \leq C\}.$$

We claim those results extend to non-orientable surfaces.

THEOREM 5.15. *Let S be a compact surface with n punctures. We exclude a sphere with $n \leq 4$, a torus with $n \leq 1$, a projective plane with $n \leq 2$ and a Klein bottle with $n \leq 1$. Then an element $g \in MCG(S)$ acts as a hyperbolic isometry on the curve complex X if and only if g is pseudo-Anosov. The action of $MCG(S)$ on X is WPD.*

PROOF. For the first part, the argument in [MM1] uses the theory of measured lamination on S . The theory does not require orientability of S , [Th].

For the second part, the argument in [BeFu] also uses measured laminations, and does not assume orientability of S . \square

We remark that Bowditch [Bo2] showed that the action of $MCG(S)$ on X is “acylindrical”, which implies WPD. His argument uses “tight geodesics” on X , which is developed in [MM2] in the orientable case.

ACKNOWLEDGEMENT.

The second author would like to thank Mustafa Korkmaz for useful discussions.

REFERENCES

- [Ba] C. Bavard, *Longueur stable des commutateurs*, Enseign. Math. (2) **37** (1991), 109–150.
- [BeFe] M. Bestvina and M. Feighn, in preparation.
- [BeFu] M. Bestvina and K. Fujiwara, *Bounded cohomology of subgroups of mapping class groups*, Geom. Topol. **6** (2002), 69–89.
- [Bo1] B. Bowditch, *Intersection numbers and the hyperbolicity of the curve complex*, J. Reine Angew. Math. **598** (2006), 105–129.
- [Bo2] B. Bowditch, *Tight geodesics in the curve complex*, preprint.
- [DS] M. J. Dunwoody and M. E. Sageev, *JSJ-splittings for finitely presented groups over slender groups*, Invent. Math. **135** (1999), 25–44.
- [FLM] B. Farb, A. Lubotzky and Y. Minsky, *Rank-1 phenomena for mapping class groups*, Duke Math. J. **106** (2001), 581–597.
- [FuP] K. Fujiwara and P. Papasoglu, *JSJ-decompositions of finitely presented groups and complexes of groups*, Geom. Funct. Anal. **16** (2006), 70–125.
- [Fu] K. Fujiwara, *The second bounded cohomology of a group acting on a Gromov-hyperbolic space*, Proc. London Math. Soc. (3) **76** (1998), 70–94.
- [Ha] W. J. Harvey, *Boundary structure of the modular group*, in: Proceedings of the 1978 Stony Brook Conference (State Univ. New York, Stony Brook, N.Y., 1978), Ann. of Math. Stud. **97**, Princeton Univ. Press, Princeton, 1981, 245–251.
- [Har] J. Harer, *Stability of the homology of the mapping class groups of orientable surfaces*, Ann. of Math. (2) **121** (1985), 215–249.

- [Iv] N. V. Ivanov, *Complexes of curves and Teichmüller modular groups*, Uspekhi Mat. Nauk **42** (1987), 49–91 (English translation: Russian Math Surveys **42** (1987), 55–107).
- [KM1] O. Kharlampovich and A. Myasnikov, *Irreducible affine varieties over a free group. I. Irreducibility of quadratic equations and Nullstellensatz*, J. Algebra **200** (1998), 472–516.
- [KM2] O. Kharlampovich and A. Myasnikov, *Irreducible affine varieties over a free group. II. Systems in triangular quasi-quadratic form and description of residually free groups*, J. Algebra **200** (1998), 517–570.
- [Ko] M. Korkmaz, *Mapping class groups of nonorientable surfaces*, Geom. Dedicata **89** (2002), 109–133.
- [L] W. B. R. Lickorish, *A finite set of generators for the homeotopy group of a 2-manifold*, Proc. Cambridge Philos. Soc. **60** (1964), 769–778.
- [MM1] H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. I. Hyperbolicity*, Invent. Math. **138** (1999), 103–149.
- [MM2] H. A. Masur and Y. N. Minsky, *Geometry of the complex of curves. II. Hierarchical structure*, Geom. Funct. Anal. **10** (2000), 902–974.
- [MS] H. Masur and J. Smillie, *Quadratic differentials with prescribed singularities and pseudo-Anosov diffeomorphisms*, Comment. Math. Helv. **68** (1993), 289–307.
- [RS] K. Rafi and S. Schleimer, *Covers and curve complex*, arXiv, math.GT/0701719.
- [Se] Z. Sela, *Diophantine geometry over groups. I. Makanin-Razborov diagrams*, Publ. Math. Inst. Hautes Études Sci. **93** (2001), 31–105.
- [Ser] C. Series, *The modular surface and continued fractions*, J. London Math. Soc. (2) **31** (1985), 69–80.
- [Stu] M. Stukow, *Dehn twists on nonorientable surfaces*, arXiv, math.GT/0601137.
- [Th] W. P. Thurston, *On the geometry and dynamics of diffeomorphisms of surfaces*, Bull. Amer. Math. Soc. (N.S.) **19** (1988), 417–431.

M. Bestvina
 Department of Mathematics
 University of Utah
 155 South 1400 East, JWB 233
 Salt Lake City, Utah 84112-0090
 USA
E-mail: bestvina@math.utah.edu

K. Fujiwara
 Graduate School of Information Science
 Tohoku University
 Sendai, 980-8579
 Japan
E-mail: fujiwara@math.is.tohoku.ac.jp

Received: 11.10.2006.